

УДК 004.932

Д.Н. Ушарова

Метод сопровождения объектов группой видеокамер с автоматической коррекцией значений матрицы проективного преобразования

Приводится описание метода сопровождения объектов группой неоткалиброванных видеокамер, позволяющего сопоставлять представления одного и того же объекта на изображениях разных камер в условиях частичных перекрытий. Метод выполняет автоматическую калибровку системы видеонаблюдения по четырем точкам с известными координатами и объекту, перемещающемуся из поля зрения одной камеры в поле зрения другой. Метод основан на коррекции значений матрицы проективного преобразования, которое вычисляется в момент нахождения объекта в пересечении полей зрения двух камер. Приведены результаты моделирования разработанного метода и получены зависимости расстояния между фактическим и прогнозируемым положением точки от площади пересечения полей зрения видеокамер.

Ключевые слова: трекинг, сопровождение объектов, видеонаблюдение, калибровка видеокамер, проективное преобразование.

doi: 10.21293/1818-0442-2016-19-3-85-89

Одной из центральных задач обработки видеоинформации является обнаружение движущегося объекта, оценка его траектории и сопоставление всех представлений одного и того же объекта с разных видеокамер. Методы сопровождения нескольких объектов несколькими видеокамерами редко рассматриваются как целостная система, большинство современных методов не используют никакой дополнительной информации от остальных видеокамер и основаны на алгоритмах сопровождения объектов одной видеокамерой. В алгоритмах сопровождения объектов одной видеокамерой задача сопровождения сводится к решению проблемы поиска соответствия объектов от одного кадра к другому. В алгоритмах сопровождения объектов несколькими видеокамерами основной задачей является решение проблемы соответствия между траекториями объектов, видимых с разных видеокамер в один момент времени.

Подходы к сопровождению объектов несколькими видеокамерами можно разделить на три основные категории: сопоставление особенностей объектов, использование информации о наблюдаемой сцене и восстановление геометрических преобразований между камерами. К первой категории относится подход, описанный в работе [1], где используют сопоставление цвета объектов. Вторая категория методов требует наличия информации о наблюдаемой сцене и точной калибровки всех камер системы [2–4]. Поиск всех представлений объекта вычисляется путем проецирования местоположения каждого объекта в мировой системе координат на карту наблюдаемой сцены. Этот подход используется в [5]. Такой метод подходит для управляемых условий, например футбольные стадионы, для которых система видеосопровождения разрабатывается специально, специально калибруются камеры и составляются точные карты наблюдаемой сцены. К третьей категории относится подход, описанный в [6], где перебирают все траектории со всех камер, сопостав-

ляют друг с другом, для каждой пары вычисляют матрицу гомографии и выполняют статистическим методом уточнение значений матрицы гомографии. С точки зрения вычислительной сложности этот метод является чрезвычайно затратным, но за большое количество итераций позволяет вычислить правильное положение объекта.

Большую популярность приобретают гибридные методы, основанные на использовании одновременно двух подходов [7]. Как правило, они включают в себя либо предварительную калибровку видеокамер, либо автоматическую в процессе сопровождения объектов, например калибровка по одному перемещающемуся по сцене объекту [8].

Постановка задачи

Недостатками вышеизложенных подходов являются: высокие требования к месту и способу установки камер, высокие требования к качеству получаемого изображения, большая вычислительная сложность, а также активное вовлечение оператора в процесс работы системы сопровождения. Целью данной работы является снижение требований к внутренним параметрам видеокамер, минимизация участия оператора и повышение быстродействия в системах видеослежения до уровня, пригодного к практическому применению.

Одной из самых требовательных к участию оператора и быстродействию процедур в задачах видеослежения является калибровка видеокамер. Обычно требуется для каждой видеокамеры указать набор точек сцены, их координаты, а также соответствующие им точки на изображениях камер. В данной работе представлен алгоритм сопровождения объектов, реализующий метод калибровки всей системы видеокамер по четырем точкам из области пересечения полей зрения двух любых из них, координаты которых задает оператор. Для всех остальных камер системы калибровка выполняется автоматически. Обязательным условием для автоматической калибровки камер является наличие пересечений

полей зрения с уже откалиброванными камерами. В случаях, когда система видеонаблюдения представляет собой группы камер, не имеющих пересечения полей зрения с другими камерами, калибровку с участием оператора необходимо выполнять для каждой группы. В таком случае задача сопоставления представления одного и того же объекта в видеокameraх из разных групп должна решаться другими методами (например, учитывать время перехода между группами, наиболее вероятный маршрут движения, характеристики объекта и т.п.).

Для сопровождения объектов группой видеокamera должен быть предварительно реализован детектор объектов одной камерой (любым методом). Все ошибки детектора объектов отразятся на работе алгоритма многокамерного сопровождения объектов, но в процессе работы алгоритма влияние этих ошибок сведется к минимуму за счет использования дополнительной информации из других камер (например, если в одной камере наблюдается перекрытие сопровождаемого объекта, он может быть виден в других камерах).

Описание предлагаемого метода калибровки

Предположим, что система состоит из множества $C = \{C_1, C_2, \dots, C_n\}$ из n камер, где каждая камера C_i имеет пересечение полей зрения хотя бы с одной камерой C_j . Для предварительной калибровки группы видеокamera предлагается использовать 4 контрольные точки с известными координатами, расположенные по сторонам прямоугольника, который находится на плоскости земли в пересечении полей зрения двух видеокamera, как изображено на рис. 1.

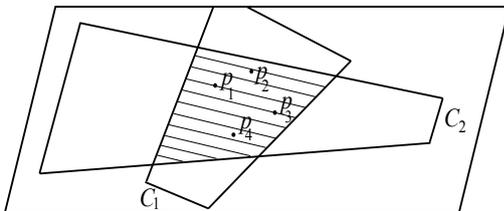


Рис. 1. Выбор контрольных точек для калибровки камер

Камеры с пересекающимися полями зрения выбираются произвольным образом, однако если прямоугольник, образуемый контрольными точками, будет занимать не менее 25% площади их изображений и располагаться ближе к центру изображений, процедура предварительной калибровки выполняется значительно точнее и быстрее. В результате ручной калибровки имеется три набора четырех точек с известными координатами:

$$P = \{P_1, P_2, P_3, P_4\},$$

$$P^1 = \{P_1^1, P_2^1, P_3^1, P_4^1\},$$

$$P^2 = \{P_1^2, P_2^2, P_3^2, P_4^2\}.$$

По ним можно найти матрицы гомографии \mathbf{H}_{gsc}^1 – из координат камеры C_1 в систему мировых

координат и \mathbf{H}_{gsc}^2 – из координат камеры C_2 в систему мировых координат.

Для этого для каждой пары соответствующих точек составляется система уравнений:

$$(u_{pk}, v_{pk}, 1)^T = \mathbf{H}(x_{pk}, y_{pk}, 1)^T,$$

где u_{pk}, v_{pk} – координаты по оси x и y точки p_k в системе координат камеры; x_{pk}, y_{pk} – координаты по оси x и y точки p_k в системе мировых координат; $k = 1, 2, 3, 4$.

Так как известна матрица гомографии \mathbf{H}_{gsc}^2 из координат камеры C_2 в систему мировых координат, выбрав из пересечения полей зрения камер C_2 и C_3 четыре точки (где каждые 3 из которых не лежат на одной прямой), можно вычислить матрицы гомографии \mathbf{H}_{gsc}^2 и \mathbf{H}_{gsc}^3 . Выполняя аналогичную процедуру, можно построить матрицы гомографии для всех камер, имеющих пересечение полей зрения, как показано на рис. 2.

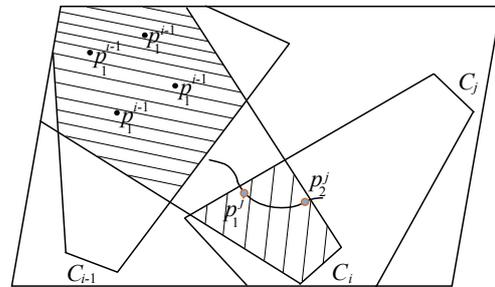


Рис. 2. Выбор точек в области пересечения полей зрения

Для того чтобы найти точки, по которым будет происходить автоматическая калибровка камер, необходимо определить границы пересечений полей зрения камер. Для вычисления пересечений полей зрения один человек свободно перемещается по сцене. Допустим, матрица гомографии \mathbf{H}_{gsc}^i известна, камеры C_i и C_j имеют пересечение полей зрения, объект находится в поле зрения камеры C_i . Ниже представлен алгоритм поиска опорных точек для вычисления матрицы гомографии \mathbf{H}_{gsc}^j .

Шаг 1. Как только объект вошел в зону пересечения полей зрения, детектор движения камеры C_j сообщает координаты его местонахождения в системе координат камеры C_j :

$$p_1^j = (u_{p1}^j, v_{p1}^j).$$

По известной матрице гомографии \mathbf{H}_{gsc}^i и известным координатам точки $p_1^i = (u_{p1}^i, v_{p1}^i)$ вычисляются координаты точки p_1 в мировой системе координат.

Как только объект покидает зону пересечения полей зрения, детектор движения камеры C_i сообщает координаты его местоположения в системе координат камеры C_i :

$$p_2^i = (u_{p2}^i, v_{p2}^i).$$

По известной матрице гомографии \mathbf{H}_{gsc}^i и известным координатам точки $p_2^i = (u_{p_2}^i, v_{p_2}^i)$ вычисляются координаты точки p_2 в мировой системе координат.

Шаг 2. Подождать до тех пор, пока объект войдет в зону пересечения полей зрения второй раз. Аналогичным образом вычислить координаты точек p_3 и p_4 .

Шаг 3. Проверить, выполняется ли условие, что любые три точки из p_1, p_2, p_3, p_4 не лежат на одной прямой (с помощью уравнения прямой, проходящей через две точки). Если условие выполняется, перейти на Шаг 4, иначе повторить Шаг 2.

Шаг 4. Вычислить матрицы гомографии \mathbf{H}_{gsc}^j и \mathbf{H}_j^i по точкам p_1, p_2, p_3, p_4 .

Уточнение проективного преобразования

После того как процедура калибровки выполнена, координаты сопровождаемого одной камерой объекта вычисляются по известной матрице гомографии и сохраняются в карте перемещений.

Как только объект попадает в область пересечения полей зрения двух камер, необходимо вычислить точки m_i и m_j в общих мировых координатах, соответствующие положению объекта на изображениях камер C_i и C_j . В общем случае $m_i \neq m_j$ (из-за погрешностей определения местоположения объекта детектором движения и погрешностей вычисления матриц гомографии).

До тех пор, пока не накопится 5 пар точек m , для уточнения значений элементов матрицы гомографии используется пять точек m, q_1, q_2, q_3, q_4 . Чтобы вычислить проективное преобразование по пяти точкам, преобразуем СЛАУ к стандартной форме записи:

$$\mathbf{A} = \begin{bmatrix} \sum x_i^2 & \sum x_i y_i & \sum u_i & 0 & 0 & 0 & -\sum u_i^2 x_i & -\sum u_i v_i x_i \\ \sum x_i y_i & \sum y_i^2 & \sum v_i & 0 & 0 & 0 & -\sum u_i v_i x_i & -\sum v_i^2 x_i \\ \sum x_i & \sum y_i & m & 0 & 0 & 0 & -\sum u_i x_i & -\sum v_i x_i \\ 0 & 0 & 0 & \sum u_i^2 & \sum u_i v_i & \sum u_i & -\sum u_i^2 y_i & -\sum u_i v_i y_i \\ 0 & 0 & 0 & \sum u_i v_i & \sum v_i^2 & \sum v_i & -\sum u_i v_i y_i & -\sum v_i^2 y_i \\ 0 & 0 & 0 & \sum u_i & \sum v_i & m & -\sum u_i y_i & -\sum v_i y_i \\ -\sum x_i^2 x'_i & -\sum x_i y_i x'_i & -\sum u_i x_i & -\sum u_i^2 y_i & -\sum u_i v_i y_i & -\sum u_i y_i & -\sum (x_i^2 + y_i^2) u_i^2 & -\sum (x_i^2 + y_i^2) u_i v_i \\ -\sum x_i y_i x'_i & -\sum y_i^2 x'_i & -\sum v_i x_i & -\sum u_i v_i y_i & -\sum v_i^2 y_i & -\sum v_i y_i & -\sum (x_i^2 + y_i^2) u_i v_i & -\sum (x_i^2 + y_i^2) v_i^2 \end{bmatrix};$$

$$\mathbf{B} = \begin{bmatrix} \sum u_i x_i \\ \sum v_i x_i \\ \sum x_i \\ \sum u_i y_i \\ \sum v_i y_i \\ \sum y_i \\ -\sum (x_i^2 + y_i^2) u_i^2 \\ -\sum (x_i^2 + y_i^2) v_i^2 \end{bmatrix};$$

$$\mathbf{A} \cdot \mathbf{G}_H = \mathbf{B}, \tag{1}$$

где \mathbf{A} – основная матрица системы; $\mathbf{G}_H = (h_{11}, h_{12}, h_{13}, h_{21}, h_{22}, h_{23}, h_{31}, h_{32})^T$ – неизвестные элементы матрицы гомографии; $\mathbf{B} = (u_1, u_2, u_3, u_4, v_1, v_2, v_3, v_4)^T$ – вектор-столбец координат ключевых точек на первом изображении.

Основная матрица системы имеет следующий вид:

$$\mathbf{A} = \begin{bmatrix} x_1 & y_1 & 1 & 0 & 0 & 0 & -u_1 x_1 & -u_1 y_1 \\ x_2 & y_2 & 1 & 0 & 0 & 0 & -u_2 x_2 & -u_2 y_2 \\ x_3 & y_3 & 1 & 0 & 0 & 0 & -u_3 x_3 & -u_3 y_3 \\ x_4 & y_4 & 1 & 0 & 0 & 0 & -u_4 x_4 & -u_4 y_4 \\ 0 & 0 & 0 & x_1 & y_1 & 1 & -v_1 x_1 & -v_1 y_1 \\ 0 & 0 & 0 & x_2 & y_2 & 1 & -v_2 x_2 & -v_2 y_2 \\ 0 & 0 & 0 & x_3 & y_3 & 1 & -v_3 x_3 & -v_3 y_3 \\ 0 & 0 & 0 & x_4 & y_4 & 1 & -v_4 x_4 & -v_4 y_4 \end{bmatrix}.$$

Перейдем к нормальной СЛАУ с помощью умножения слева обеих частей выражения (1) на матрицу \mathbf{A}^T :

$$(\mathbf{A}^T \cdot \mathbf{A}) \cdot \mathbf{G}_H = \mathbf{A}^T \cdot \mathbf{B}.$$

Для решения СЛАУ по 5 точкам добавим условие: преобразование \mathbf{H} минимизирует евклидово расстояние между точкой $\mathbf{H} \cdot \mathbf{x}$ и соответствующей ей точкой \mathbf{u} :

$$\frac{\partial F(\mathbf{G}_H)}{\partial \mathbf{G}_H} = 0,$$

где

$$\begin{aligned} \partial F(\mathbf{G}_H) = \sum_{i=1}^k & \left((u_i - (x_i h_{11} + y_i h_{12} + h_{13} - u_i x_i h_{31} - u_i y_i h_{32}))^2 + \right. \\ & \left. + (v_i - (h_{21} x_i + h_{22} y_i + h_{23} - h_{31} v_i x_i - h_{32} v_i y_i))^2 \right). \end{aligned}$$

Тогда элементы матрицы гомографии находятся из (1), как:

$$\mathbf{G}_H = (h_{11}, h_{12}, h_{13}, h_{21}, h_{22}, h_{23}, h_{31}, h_{32})^T.$$

Используя вновь найденную матрицу \mathbf{H}_{gsc}^j , необходимо пересчитать координаты точки m_j , соответствующей положению объекта на изображении камеры C_j .

Уточнение матрицы \mathbf{H}_{gsc}^j необходимо проводить до тех пор, пока евклидово расстояние D между точками m_i и m_j в общих мировых координатах вида сверху, соответствующие положению объекта

на изображениях камер C_i и C_j , не станет менее или равным требуемой погрешности оценки положения объекта.

Результаты моделирования

Для проверки разработанного метода было проведено исследование зависимости точности определения координат по матрице гомографии от площади пересечения полей зрения камер.

В ходе эксперимента поля зрения камер менялись, для того чтобы произвести вычисление матрицы гомографии с различными значениями площади пересечения полей зрения. На рис. 3 приведены примеры изображений с камер C_1, C_2, C_3, C_4 соответственно, а также наблюдаемая сцена.

Для выполнения эксперимента было сгенерировано 32 изображения (по 8 на каждую камеру) и 10 траекторий в координатах сцены. По двум траекториям и 4 точкам сцены с известными координатами выполнялась калибровка системы из 2 видеокамер предложенным методом и методом RANSAC [9,10]. Так как важным критерием работы представленных алгоритмов является быстрдействие, количество запусков методов RANSAC и процедуры уточнения проективного преобразования было ограничено до 10. Далее для каждой точки из 8 оставшихся траекторий производилось вычисление прогнозируемого положения точки на сцене по ее координатам в изображении камеры и сравнение с фактическим положением на сцене.

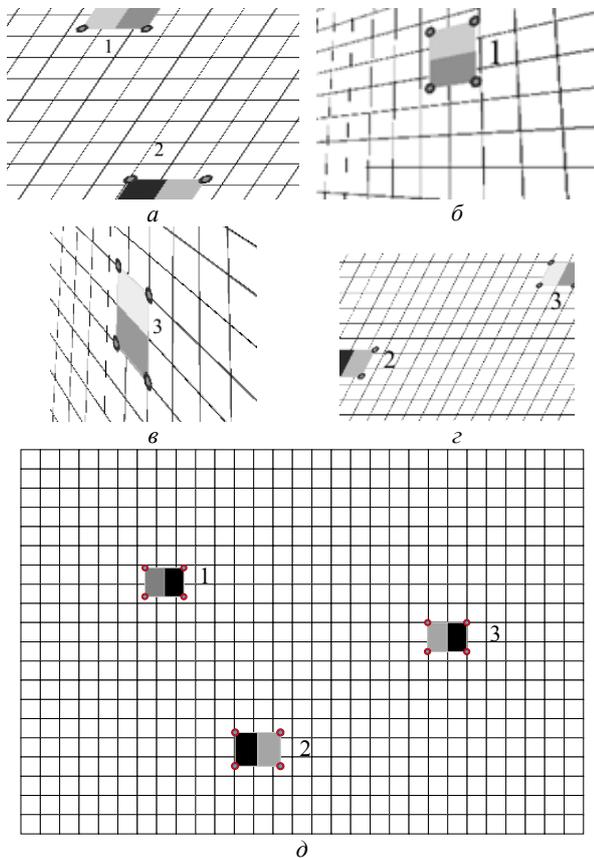


Рис. 3. Изображения: а – с камеры C_1 ; б – с камеры C_2 ; в – с камеры C_3 ; г – с камеры C_4 ; д – наблюдаемая сцена

На рис. 4 и 5 приведена зависимость среднеквадратического отклонения между фактическим и прогнозируемым положением точки от площади пересечения полей зрения видеокамер, для случая вычисления проективного преобразования по 4 точкам сцены с известными координатами.

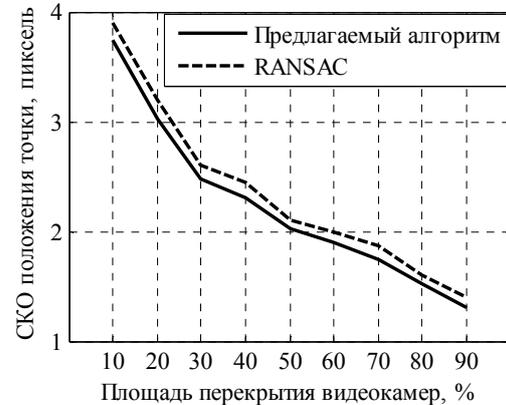


Рис. 4. SKO между фактическим и прогнозируемым положением точки для камеры C_1

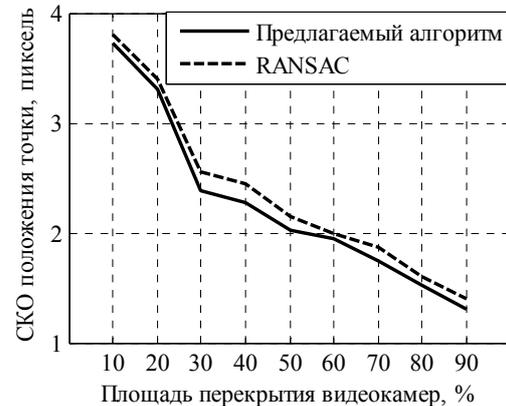


Рис. 5. SKO между фактическим и прогнозируемым положением точки для камеры C_2

Далее для калибровки камеры C_3 использовалась вычисленная на предыдущем этапе матрица гомографии H_{gsc}^2 , из пересечения полей зрения камер C_2 и C_3 находились четыре точки, по которым вычислялась матрица гомографии H_{gsc}^3 . По матрице H_{gsc}^3 и двум траекториям выполнялась калибровка камер (до 10 запусков). Аналогичным способом выполнялась калибровка камеры C_4 . На рис. 6 и 7 приведена зависимость среднеквадратического отклонения между фактическим и прогнозируемым положением точки от площади пересечения полей зрения видеокамер C_3 и C_4 .

Анализируя представленные зависимости, можно отметить, что предлагаемый метод калибровки выигрывает в точности вычисления проективного преобразования по сравнению с методом RANSAC при одинаковых условиях. В связи с тем, что в RANSAC случайным образом выбираются четверки опорных точек для вычисления проективного преобразования, более высока вероятность выбора не-

удачных пар точек [10]. Несмотря на то, что за большое количество итераций RANSAC позволяет вычислить правильное положение точки на сцене, для практического применения он не подходит.

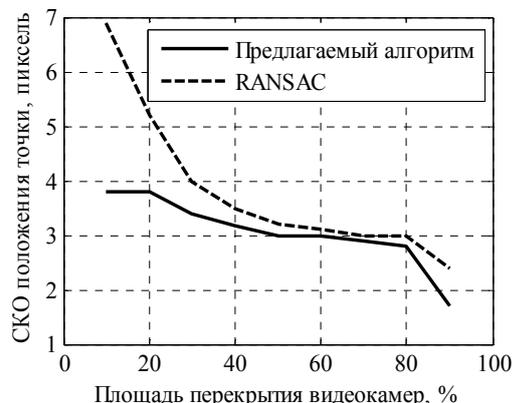


Рис. 6. СКО между фактическим и прогнозируемым положением точки для камеры C_3

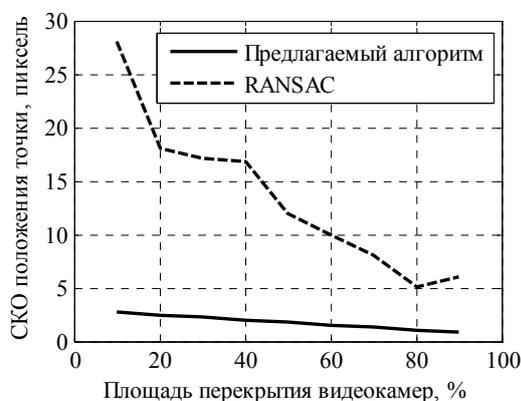


Рис. 7. СКО между фактическим и прогнозируемым положением точки для камеры C_4

Заключение

В работе приведено описание подхода к сопровождению нескольких объектов группой видеокамер. Полученные результаты свидетельствуют о возможности практического использования алгоритма в условиях с частичными или полными перекрытиями объектов. Предложенный метод калибровки видеокамер минимизирует участие оператора в работе системы, обладает необходимой точностью для распознавания всех представлений одного и того же объекта и может работать в режиме реального времени.

Литература

1. Utsumi A. Multiple-camera-based human tracking using non-synchronous observations / A. Utsumi, J. Ohya // Proc. Asian Conf. Computer Vision. – 2000. – P. 1034–1039.
2. Liebowitz D. Metric rectification for perspective images of planes / D. Liebowitz, A. Zisserman // Computer Vision and Pattern Recognition, 1998. Proceedings. 1998 IEEE

Computer Society Conference on. – IEEE, 1998. – P. 482–488.

3. Hemayed E. E. A survey of camera self-calibration // Advanced Video and Signal Based Surveillance, 2003. Proceedings. IEEE Conference on. – IEEE, 2003. – P. 351–357.

4. Kahl F. Critical motions for auto-calibration when some intrinsic parameters can vary / F. Kahl, B. Triggs, K. Åström // Journal of Mathematical Imaging and Vision. – 2000. – Vol. 13, No. 2. – P. 131–146.

5. An architecture for multiple perspective interactive video / P. Kelly, A. Katkere, D. Kuramura, S. Moezzi, S. Chatterjee, R. Jain // Proceedings of the third ACM international conference on Multimedia. – ACM, 1995. – P. 201–212.

6. Lee L. Monitoring activities from multiple video streams: Establishing a common coordinate frame / L. Lee, R. Romano, G. Stein // IEEE Transactions on pattern analysis and machine intelligence. – 2000. – Vol. 22, No. 8. – P. 758–767.

7. Tracking many objects with many sensors / H. Pasula, S. Russell, M. Ostland, Ya. Ritov // IJCAI. – 1999. – Vol. 99. – P. 1160–1171.

8. Hartley R.I. An algorithm for self calibration from several views // CVPR. – 1994. – Vol. 94. – P. 908–912.

9. Fischler M.A. Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography / M.A. Fischler, R.C. Bolles // Communications of the ACM. – 1981. – Vol. 24, No. 6. – P. 381–395.

10. Liu Z. A New Efficient Calibration Method for Binocular Camera / Z. Liu, J. Fang, B. Zhou // International Journal of Signal Processing, Image Processing and Pattern Recognition. – 2016. – Vol. 9. – No. 2. – P. 445–456.

Ушарова Дарья Николаевна

Аспирантка каф. телекоммуникаций и основ радиотехники (ТОР) Томского университета систем управления и радиоэлектроники (ТУСУР)
Тел.: +7-913-111-66-30
Эл. почта: udn@csp.tusur.ru

Usharova D.N.

Method for objects tracking using a group of cameras with automatic correction of a projective transformation matrix

A description of a method for tracking objects by a group of uncalibrated cameras, allowing to compare the representations of the same object in the images of different cameras in partial overlap is described. The method automatically calibrates the video surveillance system on four points with known coordinates of the object and moving out from the field of view of one camera to field of view of another. The method is based on the correction of the projective transformation matrix, which is calculated at the moment when the object appear in the fields of view of the two cameras. The results of simulation are presented and distance relations of actual and predicted positions to the surface of intersection of camera views are obtained.

Keywords: tracking, object tracking, video surveillance, cameras calibration, projective transformation.